

09/878,032

**REMARKS**

In view of the following discussion, the Applicant submits that none of the claims now pending in the application is anticipated under the provisions of 35 U.S.C. § 102 or made obvious under the provisions of 35 U.S.C. § 103. Thus, the Applicant believes that all of these claims are now in allowable form.

**I. REJECTION OF CLAIMS 1-5, 17-21, 25-29, 33 AND 34 UNDER 35 U.S.C. § 102**

The Examiner has rejected claims 1-5, 17-21, 25-29, 33 and 34 in the Office Action as being anticipated by the Ellozy et al. patent (U.S. patent 5,649,060, issued on July 15, 1997, hereinafter Ellozy). In response, the Applicant has amended independent claims 1, 17, 25, 33 and 34 from which claims 2-5, 18-21, 26-29 depend, in order to more clearly recite aspects of the invention.

Ellozy discloses a method of aligning a written transcript comprising index words with speech in audio clips. Ellozy discloses using an automatic speech recognizer to decode the speech in the audio clip and produce a text file comprising recognized words having a high probability of corresponding to acoustic information signal units in the audio clip. Each of the recognized words is time stamped by a time aligner while the audio data is decoded. A comparator matches the recognized words with index words from the transcript via identification of similar words or clusters of words, and tags each matched index word with the time stamp of the recognized word. Thus, an index word is associated with a time stamp through its match with a recognized word. An index word that is not paired to a recognized word is tagged with a recording location based on an interpolation between the time stamps of recognized words nearest to the context of the unpaired index word. Ellozy does not teach, show or suggest, however, that the time aligner can time stamp the index words with recording locations of the audio clip.

The Examiner's attention is directed to the fact that Ellozy fails to disclose or suggest the novel method of constructing a digital talking book wherein a plurality of time points of the audio data are linked to a plurality of synchronizable elements of the text data which are produced independently of the plurality of time point of the audio

09/878,032

data, as claimed in Applicant's independent claims 1, 17, 25, 33 and 34. Specifically, Applicant's claims 1, 17, 25, 33 and 34, as amended, positively recite:

1. Method for constructing a digital talking book from text data and audio data, said method comprising the steps of:

- (a) accessing a first synchronization file that identifies a plurality of synchronizable elements of the text data;
- (b) accessing a second synchronization file that identifies a plurality of time points of the audio data, wherein said plurality of synchronizable elements of the text data are produced independently of said plurality of time points of the audio data; and
- (c) building links between said identified synchronizable elements of the text data with said identified time points of the audio data. (Emphasis added)

17. A computer-readable medium having stored thereon a plurality of instructions, the plurality of instructions including instructions which, when executed by a processor, cause the processor to perform the steps comprising of:

- (a) accessing a first synchronization file that identifies a plurality of synchronizable elements of the text data;
- (b) accessing a second synchronization file that identifies a plurality of time points of the audio data, wherein said plurality of synchronizable elements of the text data are produced independently of said plurality of time points of the audio data; and
- (c) building links between said identified synchronizable elements of the text data with said identified time points of the audio data. (Emphasis added)

25. Apparatus for constructing a digital talking book from text data and audio data, said apparatus comprising:

means for accessing a first synchronization file that identifies a plurality of synchronizable elements of the text data and for accessing a second synchronization file that identifies a plurality of time points of the audio data, wherein said plurality of synchronizable elements of the text data are produced independently of said plurality of time points of the audio data; and

means for building links between said identified synchronizable elements of the text data with said identified time points of the audio data. (Emphasis added)

33. A computer readable medium having stored thereon a data structure for assisting in the construction of a digital talking book from text data and audio data, said text data comprising a plurality of synchronizable elements, said audio data comprising plurality of time points, and wherein said plurality of synchronizable elements of the text data are produced independently of said plurality of time points of the audio data, said data structure comprising:

- a project metadata field;
- a project text data field; and
- a synchronizable element field. (Emphasis added)

09/878,032

34. A computer readable medium having stored thereon a data structure for assisting in the construction of a digital talking book from text data and audio data, said text data comprising a plurality of synchronizable elements, said audio data comprising plurality of time points, and wherein said plurality of synchronizable elements of the text data are produced independently of said plurality of time points of the audio data, said data structure comprising:

a data element field, wherein said data comprises at least one record element field, wherein said at least one record element field comprises:

- a identification field;
- a starttime field;
- an endtime field;
- and a type field. (Emphasis added)

Applicant's invention is directed to a method and apparatus for constructing a digital talking book from text data and audio data. To convert existing analog-recorded books into digital talking books, the method must align text data and audio data that are produced independently of each other.

The present invention provides a method and apparatus for constructing a digital talking book from text data having a plurality of synchronizable elements and audio data having a plurality of time points, wherein the plurality of synchronizable elements are produced independently of the plurality of time points. In one embodiment of the method, a first synchronization file is accessed that identifies the plurality of synchronizable elements, a second synchronization file is accessed that identifies the plurality of independently produced time points, and links are built between the plurality synchronizable elements and the plurality of independently produced time points. By building links between the plurality synchronizable elements and the plurality of independently produced time points, the method is beneficially applicable to utilizing the vast libraries of existing analog recorded books.

In contrast, Ellozy discloses a method in which words which are recognized, by using a speech recognizer, from the audio clip, are time stamped with recording locations of the audio clip. Clearly, the recognized words are generated directly from the audio clip. Thus, Ellozy fails to anticipate or make obvious Applicant's invention.

Specifically, Ellozy discloses that the recognized words are give a time stamp during the speech recognition process that decodes the recognized words from the

09/878,032

audio clip. Ellozy then tags index words which are paired with recognized words with the recognized words' time stamp, indirectly associating the index word with a recording location in the audio clip. Thus, Ellozy discloses linking recognized words with recording locations, and linking index words with recognized words, but fails to teach or make obvious building links directly between index words and recording locations, or linking synchronizable elements of the text data and independently produced time points of the audio data, as positively claimed by the Applicant in amended claims 1, 17, 25, 33 and 34. Therefore, the Applicant submits that independent claims 1, 17, 25, 33 and 34, as amended, fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

Dependent claims 2-5, 18-21, 26-29 depend from claims 1, 17 and 25, and recite additional features therefore. As such, and for at least the reasons set forth above, the Applicant submits that claims 2-5, 18-21, 26-29 are not anticipated by the teachings of Ellozy. Therefore, the Applicant submits that dependent claims 2-5, 18-21, 26-29 also fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

## **II. REJECTION OF CLAIMS 6, 7, 22, 23, 30 AND 31 UNDER 35 U.S.C. § 103**

The Examiner rejected claims 6, 7, 22, 23, 30 and 31 under 35 U.S.C. §103(a) as being unpatentable over Ellozy in view of the allegedly well-known art. The Applicant has amended independent claims 1, 17 and 25, from which claims 6, 7, 22, 23, 30 and 31 depend, as described above in order to more clearly recite aspects of the invention. The remainder of the rejection is respectfully traversed.

Ellozy has been discussed above. The Examiner, with respect to Claim 6, has taken Official Notice "on the playback audio by clicking the corresponding text data" at page 4 of Paper No. 1. The Applicant respectfully traverses the Official Notice taken by the Examiner. A method of constructing a talking book including the step "clicking on one of said synchronizable elements on said display to play said linked associated audio data," as recited in Claim 6, is not well known in the art of constructing a talking digital book. The step recited in Claim 6 allows an operator to quickly verify which synchronizable elements of the text data have been linked to time points of the audio

09/878,032

data, as well as the accuracy of links between synchronizable elements and time points, as described in paragraphs 43 and 46 of the Specification.

The Examiner, with respect to Claim 7, has also taken Official Notice "on the highlighted the text" at page 4 of Paper No. 1. The Applicant respectfully traverses the Official Notice taken by the Examiner. A method of constructing a talking book including the step "clicking on one of said synchronizable elements on said display to display said linked associated text data as being highlighted," as recited in Claim 7, is not well known in the art of constructing a talking digital book. As discussed above in regards to Claim 6, the step recited in Claim 7 allows an operator to quickly verify which synchronizable elements of the text data have been linked to time points of the audio data, and to check the accuracy of those links.

The Examiner's attention is also directed to the fact that Ellozy and the allegedly well-known art (either singly or in any permissible combination) fail to disclose or suggest the novel method of constructing a digital talking book including linking a plurality of synchronizable elements of text data with an independently produced plurality of time points of audio data, as claimed in Applicant's independent claims 1, 17 and 25. Applicant's claims 1, 17 and 25, have been recited above.

As recited in the preceding claim, Applicant's invention teaches a method and apparatus for constructing a digital talking book from text data having a plurality of synchronizable elements and audio data having a plurality of time points, wherein the plurality of synchronizable elements are produced independently of the plurality of time points. In one embodiment of the method, a first synchronization file is accessed that identifies the plurality of synchronizable elements, a second synchronization file is accessed that identifies the plurality of independently produced time points, and links are built between the plurality synchronizable elements and the plurality of independently produced time points.

In contrast, the combination of Ellozy and the allegedly well-known art at most disclose linking recognized words with recording locations in audio data from which the words are recognized, or are dependent upon. Thus, Ellozy and the allegedly well-known art, singularly and in combination, fail to anticipate or make obvious Applicants'

09/878,032

invention.

Specifically, the combination of Ellozy and the allegedly well-known art disclose that words may be recognized or generated from audio data, and that these recognized or generated words may be assigned a time stamp or recording location in the audio data from which they are recognized. The combination of Ellozy and the allegedly well-known art disclose that the recognized words may be matched to words in a text file, and thus the words in the text file are linked to recognized words, and not to time points of audio data. Furthermore, the recognized words of the combination of Ellozy and the allegedly well-known art are not matched to independently produced time points of the audio data, but rather are matched to time points in the audio data from which they were generated or are dependent upon. Thus, links are not built between synchronizable elements of text data and independently produced time points of audio data, as is recited in the Applicant's claims. Ellozy and the allegedly well-known art, singularly and in combination, thus fail to teach or make obvious a method of constructing a digital talking book wherein links are built between synchronizable elements of the text data and independently produced time points of the audio data, as positively claimed by the Applicant in amended claims 1, 17 and 25. Therefore, the Applicant submits that independent claims 1, 17 and 25, as amended, fully satisfy the requirements of 35 U.S.C. § 103 and are patentable thereunder.

Dependent claims 6, 7, 22, 23, 30 and 31 depend, either directly or indirectly, from claims 1, 17 and 25 and recite additional features thereof. As such and for at least the same reasons set forth above, the Applicant submits that claims 6, 7, 22, 23, 30 and 31 are also not made obvious by the teachings of Ellozy in view of the allegedly well-known art. Therefore, the Applicant submits that dependent claims 6, 7, 22, 23, 30 and 31 also fully satisfy the requirements of 35 U.S.C. § 103 and are patentable thereunder.

### **III. REJECTION OF CLAIMS 8-16, 24 AND 32 UNDER 35 U.S.C. § 103**

The Examiner rejected claims 8-16, 24 and 32 under 35 U.S.C. §103(a) as being unpatentable over Ellozy in view of the Van Thong et al. patent (U.S. Patent No. 6,442,518, issued August 27, 2002, hereinafter Van Thong). The Applicant has

09/878,032

amended independent claims 1, 17 and 25, from which claims 8-16, 24 and 32 depend, as described above in order to more clearly recite aspects of the invention. The remainder of the rejection is respectfully traversed.

Ellozy has been discussed above. Van Thong discloses a method and apparatus for refining time alignments of roughly aligned closed captions with audio data. The method includes performing a speech recognition operation in which a sequence of words is generated from the audio data, and these generated words are matched to words in the closed captions. The generated words and the closed caption words have associated time stamps, and the time stamps of the generated words are used to modify the time stamps of the closed captioned words to refine the alignment of the closed captions to the audio data.

The Examiner's attention is directed to the fact that Ellozy and Van Thong, either singly or in any permissible combination, fail to disclose or suggest the novel method of constructing a digital talking book from text data and audio data, wherein a plurality of synchronizable elements of the text data are linked to a plurality of independently produced time points of the audio data, as claimed in Applicant's independent claims 1, 17 and 25. Applicant's independent claims 1, 17 and 25 have been recited above.

As recited in the preceding claim, Applicant's invention teaches a method and apparatus for constructing a digital talking book in which a first synchronization file is accessed that identifies a plurality of synchronizable elements of the text data, a second synchronization file is accessed that identifies a plurality of time points of the audio data, wherein the plurality of synchronizable elements of the text data are produced independently of the plurality of time points of the audio data, and links are built between the plurality synchronizable elements and the plurality of independently produced time points.

In contrast, the combination of Ellozy and Van Thong at most discloses linking recognized words via a recognizer with recording locations in audio data from which the words are recognized, or are dependent upon. Thus, Ellozy and Van Thong, singularly and in combination, fail to anticipate or make obvious Applicant's invention.

Specifically, the combination of Ellozy and Van Thong disclose that words may

09/878,032

be recognized or generated from audio data, and that these recognized or generated words may be assigned a time stamp or recording location in the audio data from which they are recognized. The combination of Ellozy and Van Thong disclose that the recognized words may be matched to words in a text file, and thus the words in the text file are linked to recognized words, and not to time points of audio data. Furthermore, the recognized words of the combination of Ellozy and Van Thong are not matched to independently produced time points of the audio data, but rather are matched to time points in the audio data from which they were generated or are dependent upon. Thus, links are not built between synchronizable elements of text data and independently produced time points of audio data, as is recited in the Applicant's claims. Ellozy and Van Thong, singularly and in combination, thus fail to teach or make obvious a method of constructing a digital talking book wherein links are built between synchronizable elements of the text data and independently produced time points of the audio data, as positively claimed by the Applicant in amended claims 1, 17 and 25. Therefore, the Applicant submits that independent claims 1, 17 and 25, as amended, fully satisfy the requirements of 35 U.S.C. § 103 and are patentable thereunder.

Dependent claims 8-16, 24 and 32 depend, either directly or indirectly, from claims 1, 17 and 25 and recite additional features thereof. As such and for at least the same reasons set forth above, the Applicant submits that claims 8-16, 24 and 32 are also not made obvious by the teachings of Ellozy in view of Van Thong. Therefore, the Applicant submits that dependent claims 8-16, 24 and 32 also fully satisfy the requirements of 35 U.S.C. § 103 and are patentable thereunder.

#### **IV. CONCLUSION**

Thus, the Applicant submits that all of the presented claims now fully satisfy the requirements of 35 U.S.C. §102 and §103. Consequently, the Applicant believes that all of these claims are presently in condition for allowance. Accordingly, both reconsideration of this application and its swift passage to issue are earnestly solicited.

If, however, the Examiner believes that there are any unresolved issues requiring the issuance of a final action in any of the claims now pending in the application, it is




09/878,032

requested that the Examiner telephone Mr. Kin-Wah Tong, Esq. at (732) 530-9404 so that appropriate arrangements can be made for resolving such issues as expeditiously as possible.

1/27/05  
Date

Moser, Patterson & Sheridan, LLP  
595 Shrewsbury Avenue  
Shrewsbury, New Jersey 07702

Respectfully submitted,

  
Kin-Wah Tong, Attorney  
Reg. No. 39,400  
(732) 530-9404